Safe Haven Challenge

OCTOBER 2024

Bob Foreman
Software Engineer Lead
LexisNexis Risk Solutions

NUHacks2024 **Challenge!**

Safe Haven!

Many travelers find themselves on occasion in a strange land and a strange city, sometimes without their control. Without knowing the risks and dangers in that area, sometimes a tourist can suddenly be in a dangerous situation.

What can we do as developers to help prevent this?

This year's challenge will analyze different social factors by area such as poverty, unemployment, and many other factors to assess the risk for the traveler in a strange land.

The goal of the challenge is to answer two questions:

1. Analysis of social factors in an area (unemployment, education, poverty, and population) and identify it as a "Hot Spot".
2. Provide additional information to the traveler to help find "safe haven" resources in their area (fire and police stations, hospitals, churches, food banks, etc.)

HPCC SYSTEMS

The Data!

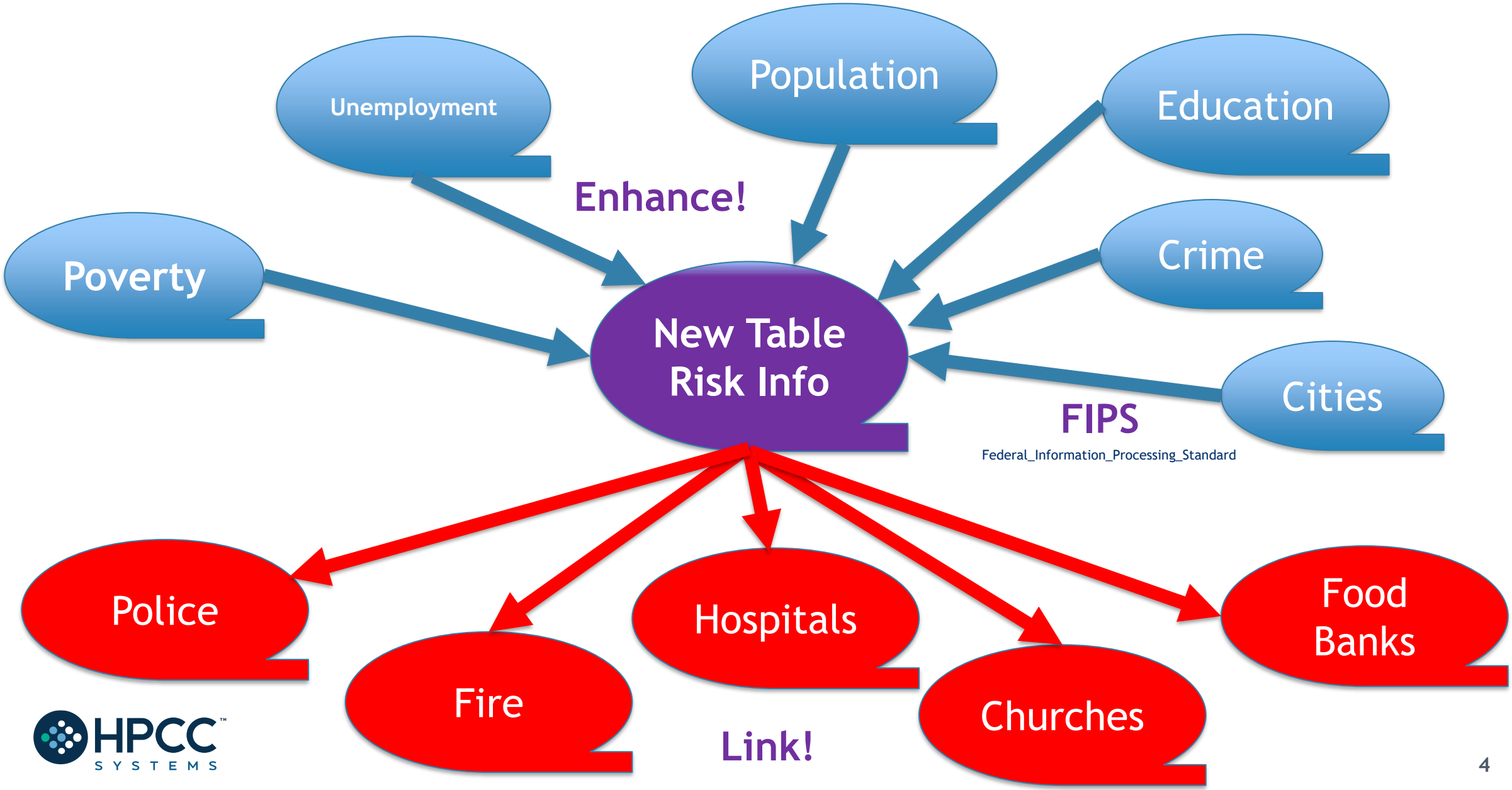
City and County Data has been collected from all 50 US states and organized into a simple dataset to use as your source. In addition, many public datasets have also been gathered and cleaned to help get you started.

These datasets include:
Education
Unemployment
Poverty
Population
Crime
Police
Fire Stations
Hospitals
Places of Worship
Food Banks

Reference Datasets:
A **Cities** dataset with related FIPS and Zip Codes (used for linking the above datasets to a new "Risk" dataset)
Unemployment Rates (Not really used in this challenge but interesting data!)
You are not limited to using these datasets! Extra credit will be rewarded by linking in other pertinent data!

HPCC SYSTEMS

Data Flow


Unemployment
Population
Education
Enhance!
Poverty
Crime
New Table
Risk Info
Cities
FIPS
Federal_Information_Processing_Standard
Police
Fire
Hospitals
Churches
Food
Banks
Link!

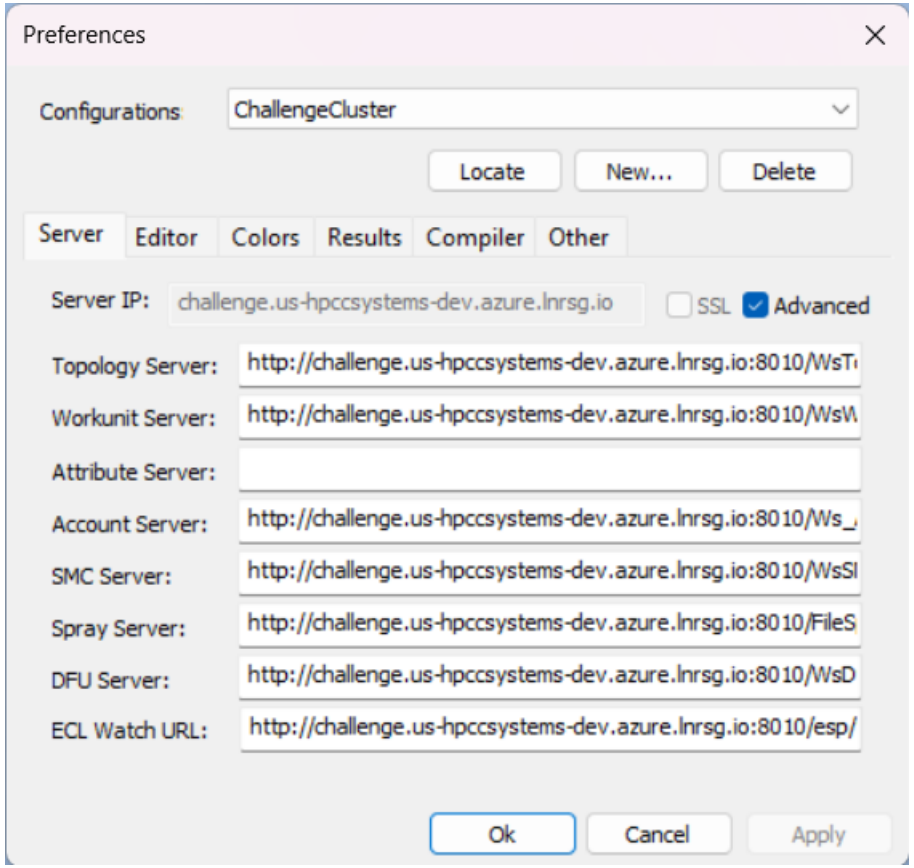
The Playing Field!

HPCC Cluster ECL Watch:

http://challenge.us-hpccsystems-dev.azure.lnrsg.io:8010

```
{
    "name": "Challenge",
    "type": "ecl",
    "request": "launch",
    "protocol": "http",
    "serverAddress": "challenge.us-hpccsystems-dev.azure.lnrsg.io",
    "port": 8010,
    "path": "",
    "targetCluster": "thor",
    "abortSubmitOnError": true,
    "rejectUnauthorized": true,
    "resultLimit": 100,
    "timeoutSecs": 60,
    "user": "Bob_F",
    "password": ""
},
```

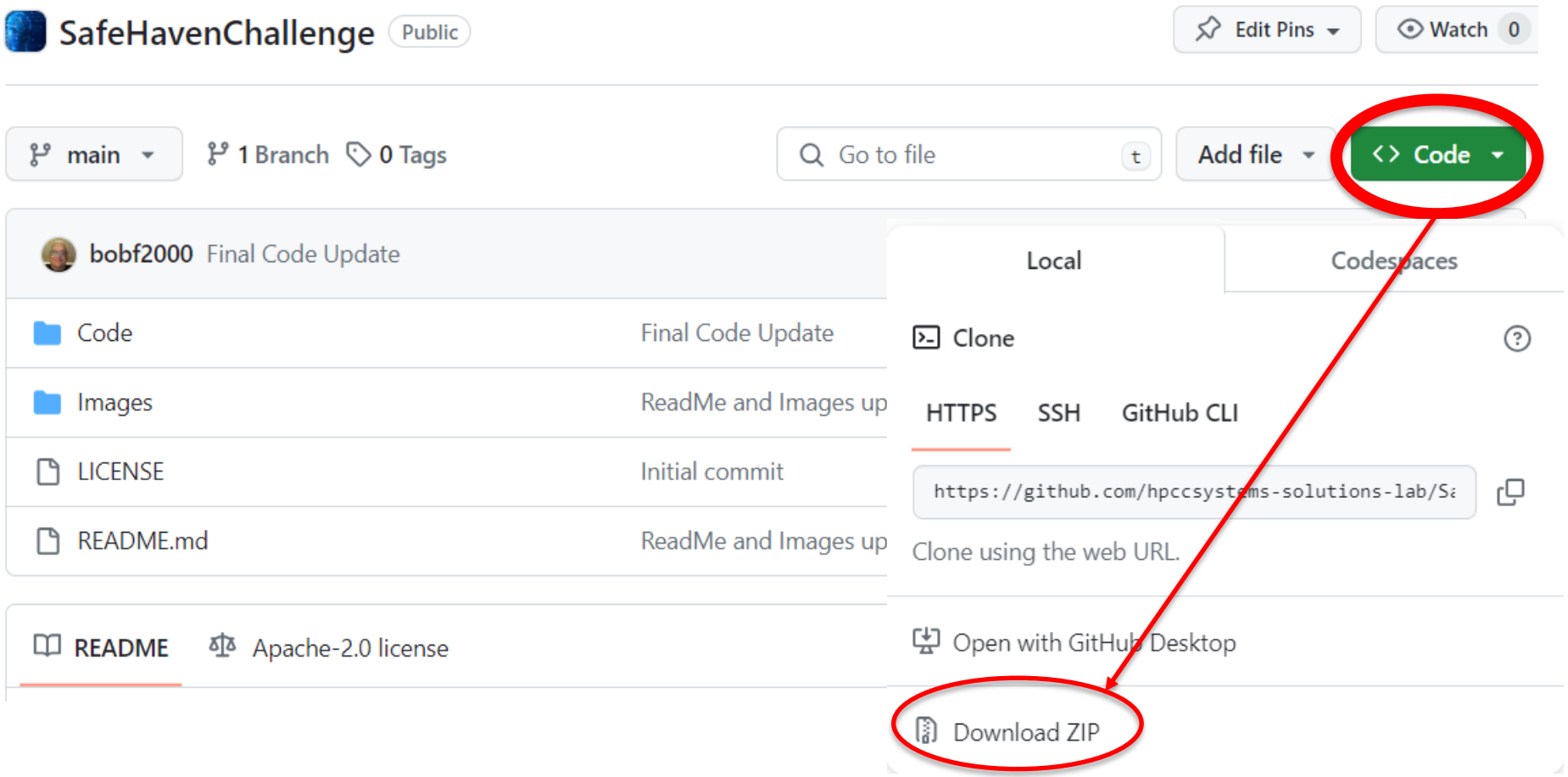
{} launch.json ×

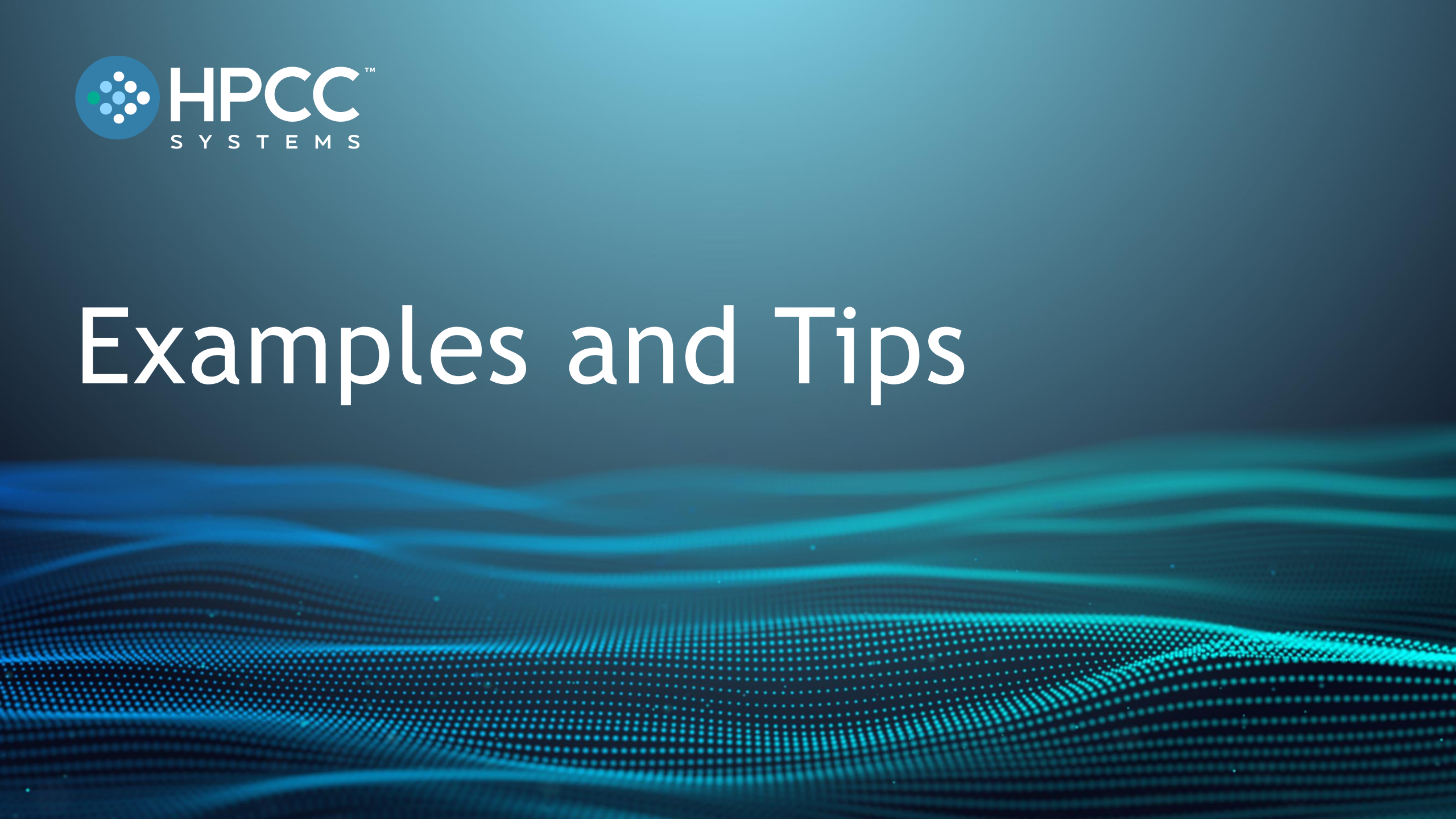


ECL IDE

The Repo!

https://github.com/hpccsystems-solutions-lab/SafeHavenChallenge





Examples and Tips

Delivering your Results!

Create a Core Dataset that aggregates your analysis:

```
// Let's create a core "risk" file that the county code (FIPS) and the primary city.
// We can extra ct this data from the Cities file.
IMPORT $;
CityDS := $.File_AllData.City_DS;
Crime  := $.File_AllData.CrimeDS;

//CityDS(county_fips = 5035); Test to verify data accuracy for the crime score


// Declare our core RECORD:
RiskRec := RECORD
    STRING45  city;
    STRING2   state_id;
    STRING20  state_name;
    UNSIGNED3 county_fips;
    STRING30  county_name;
END;

BaseInfo := PROJECT(CityDS,RiskRec);
OUTPUT(BaseInfo,NAMED('BaseData'));

RiskPlusRec := RECORD
 BaseInfo;
 EducationScore  := 0;
 PovertyScore    := 0;
 PopulationScore := 0;
 CrimeScore      := 0;
 Total           := 0;
END;

RiskTbl := TABLE(BaseInfo,RiskPlusRec);
OUTPUT(RiskTbl,NAMED('BuildTable'));
```

HPCC SYSTEMS

Delivering your Results!

Create a Core Dataset that aggregates your analysis:

```
34  //Let's add a Crime Score!
35
36  CrimeRec := RECORD
37  CrimeRate := TRUNCATE((INTEGER)Crime.crime_rate_per_100000);
38  Crime.fips_st;
39  fips_cty := (INTEGER)Crime.fips_cty;
40  Fips := Crime.fips_st + INTFORMAT(Crime.fips_cty,3,1);
41  END;
42
43  CrimeTbl := TABLE(Crime,CrimeRec);
44  OUTPUT(CrimeTbl,NAMED('BuildCrimeTable'));
45
46  JoinCrime := JOIN(CrimeTbl,RiskTbl,
47                    LEFT.fips = (STRING5)RIGHT.county_fips,
48                    TRANSFORM(RiskPlusRec,
49                              SELF.CrimeScore := LEFT.crimerate,
50                              SELF            := RIGHT),
51                              RIGHT OUTER);
52
53  OUTPUT(SORT(JoinCrime,-CrimeScore),NAMED('AddedCrimeScore'));
54
55  //Now go out and get the others! Good like with your challenge!
56  //After you complete the other scores, make sure to OUTPUT to a file and then create a DATASET so
57  //that you can reference and deliver it to the judges.
58
59
```

HPCC SYSTEMS

Delivering your Results!

Extracting Data using our STRING library:

Submit | Target: thor

```
IMPORT $,STD;

UNEMP     := $.File_AllData.unemp_byCountyDS;
EDU       := $.File_AllData.EducationDS;
POVTY     := $.File_AllData.pov_estimatesDS;


//Add Poverty Percentage ages 0-17 for FIPS area:
POVTBL := TABLE(POVTY((STD.Str.Find(attribute, 'PCTPOV017_2021',1) <> 0)),
                {Fips_Code,attribute,value});
OUTPUT(SORT(POVTBL,-value),NAMED('PovertyPct0to17'));

//Add Unemployment Rate for area:
CT_UNEMP := TABLE(UNEMP((STD.Str.Find(attribute, 'Unemployment_rate',1) <> 0)),
                {Fips_Code,cnt := ROUND(AVE(GROUP,value),2)},Fips_Code);
OUTPUT(SORT(CT_UNEMP,-cnt),NAMED('UNEMP_Rate'));

EDU_CT_FIPS := TABLE(EDU((STD.Str.Find(attribute, 'Percent of adults with less than a high school diploma',1) <> 0)),
                {Fips_Code,tot := ROUND(AVE(GROUP,value),2)},fips_code);
OUTPUT(SORT(EDU_CT_FIPS,-tot),NAMED('NoHighSch'));

```

HPCC SYSTEMS


HPCC SYSTEMS
Data Delivery
(Roxie and Visualization)

Delivering your Results!

Step 1 - Clean, Declare and Build your Indexes (Churches):

```
IMPORT $,STD;
//This file is used to demonstrate how to "clean" a raw dataset (Churches) and create an index to be used in a ROXIE service
Churches := $.File_AllData.ChurchDS;
Cities   := $.File_AllData.City_DS;


//First, determine what fields you want to clean:
CleanChurchRec := RECORD
    STRING70  name;
    STRING35  street;
    STRING22  city;
    STRING2   state;
    UNSIGNED3 zip;
    UNSIGNED1 affiliation;
    UNSIGNED3 PrimaryFIPS; //New - will be added from Cities DS
END;
//PROJECT is used to transform one data record to another.
CleanChurch := PROJECT(Churches,TRANSFORM(CleanChurchRec,
                                          SELF.name               := STD.STR.ToUpperCase(LEFT.name),
                                          SELF.street             := STD.STR.ToUpperCase(LEFT.street),
                                          SELF.city               := STD.STR.ToUpperCase(LEFT.city),
                                          SELF.State              := STD.STR.ToUpperCase(LEFT.state),
                                          SELF.zip                := LEFT.zip,
                                          SELF.affiliation        := LEFT.affiliation,
                                          SELF.PrimaryFIPS        := 0));
```

Delivering your Results!

Step 1 - Clean, Declare and Build your Indexes (Churches):

```
26  //JOIN is used to combine data from different datasets
27  CleanChurchFIPS :=       JOIN(CleanChurch,Cities,
28                            LEFT.city  = STD.STR.ToUpperCase(RIGHT.city) AND
29                            LEFT.state = RIGHT.state_id,
30                            TRANSFORM(CleanChurchRec,
31                                      SELF.PrimaryFIPS := (UNSIGNED3)RIGHT.county_fips,
32                                      SELF             := LEFT),LEFT OUTER,LOOKUP);
33  //Write out the new file and then define it using DATASET
34  WriteChurches      := OUTPUT(CleanChurchFIPS,,'~SAFE::OUT::Churches',OVERWRITE);
35  CleanChurchesDS    := DATASET('~SAFE::OUT::Churches',CleanChurchRec,FLAT);
36
37  //Declare and Build Indexes (special datasets that can be used in the ROXIE data delivery cluster
38  CleanChurchIDX     := INDEX(CleanChurchesDS,{city,state},{CleanChurchesDS},'~SAFE::IDX::Church::CityPay');
39  CleanChurchFIPSIDX := INDEX(CleanChurchesDS,{PrimaryFIPS},{CleanChurchesDS},'~SAFE::IDX::Church::FIPSPay');
40  BuildChurchIDX     := BUILD(CleanChurchIDX,OVERWRITE);
41  BuildChurchFIPSIDX := BUILD(CleanChurchFIPSIDX,OVERWRITE);
42
43  //SEQUENTIAL is similar to OUTPUT, but executes the actions in sequence instead of the default parallel actions of the HPCC
44  SEQUENTIAL(WriteChurches,BuildChurchIDX,BuildChurchFIPSIDX);
45
```

Delivering your Results!

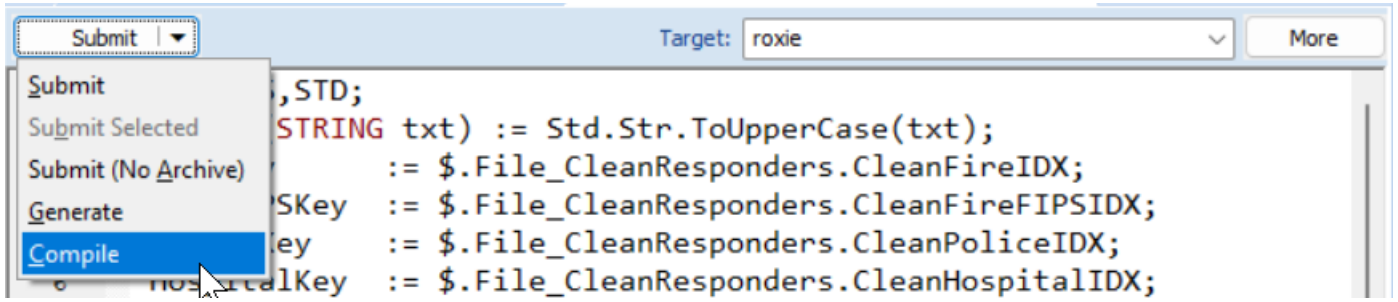
Step 2 - Design and Write Your Query:

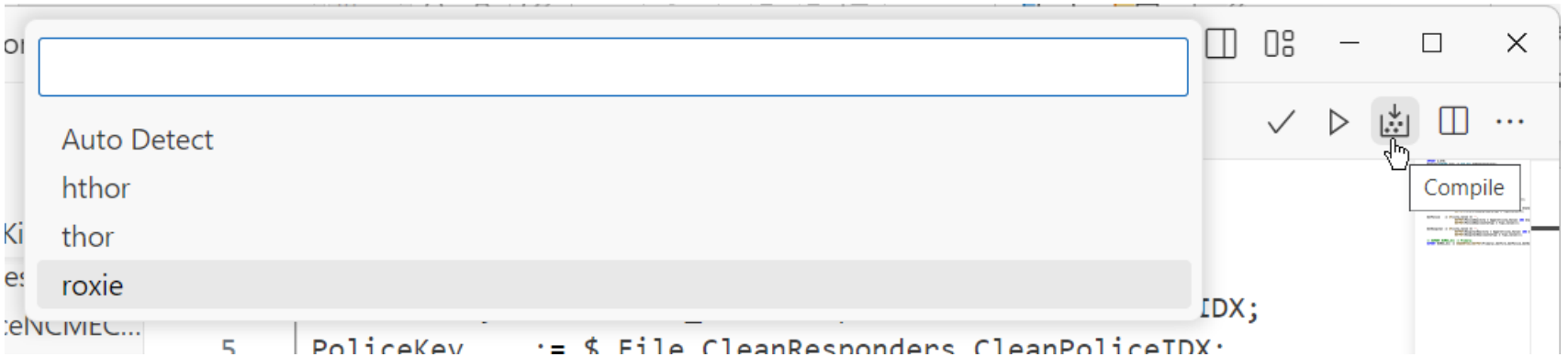
```
IMPORT $,STD;
UpperIt(STRING txt) := Std.Str.ToUpperCase(txt);
//These INDEXes are created (built) in BWR_CleanChurches
CleanChurchRec := RECORD
    STRING70  name;
    STRING35  street;
    STRING22  city;
    STRING2   state;
    UNSIGNED3 zip;
    UNSIGNED1 affiliation;
    UNSIGNED3 PrimaryFIPS; //New - will be added from Cities DS
END;
CleanChurchesDS    := DATASET('~SAFE::OUT::Churches',CleanChurchRec,FLAT);
CleanChurchIDX     := INDEX(CleanChurchesDS,{city,state},{CleanChurchesDS},'~SAFE::IDX::Church::CityPay');
CleanChurchFIPSIDX := INDEX(CleanChurchesDS,{PrimaryFIPS},{CleanChurchesDS},'~SAFE::IDX::Church::FIPSPay');

/* To Publish your Query:
   1. Change Target to ROXIE
   2. Compile ONLY
   3. Open ECL Watch and select the Publish tab to publish your query
   4. Test and demonstarte using: http://training.us-hpccsystems-dev.azure.lnrsg.io:8002

*/
EXPORT Safe_Svc(FipsVal,STRING22 CityVal,STRING2 StateVal) := FUNCTION
MyChurch := IF(FipsVal = 0,
               OUTPUT(CleanChurchIDX(City=UpperIt(CityVal),State=UpperIt(StateVal))),
               OUTPUT(CleanChurchFIPSIDX(PrimaryFIPS=FipsVal)));
RETURN MyChurch;
END;

```

HPCC SYSTEMS

Delivering your Results!

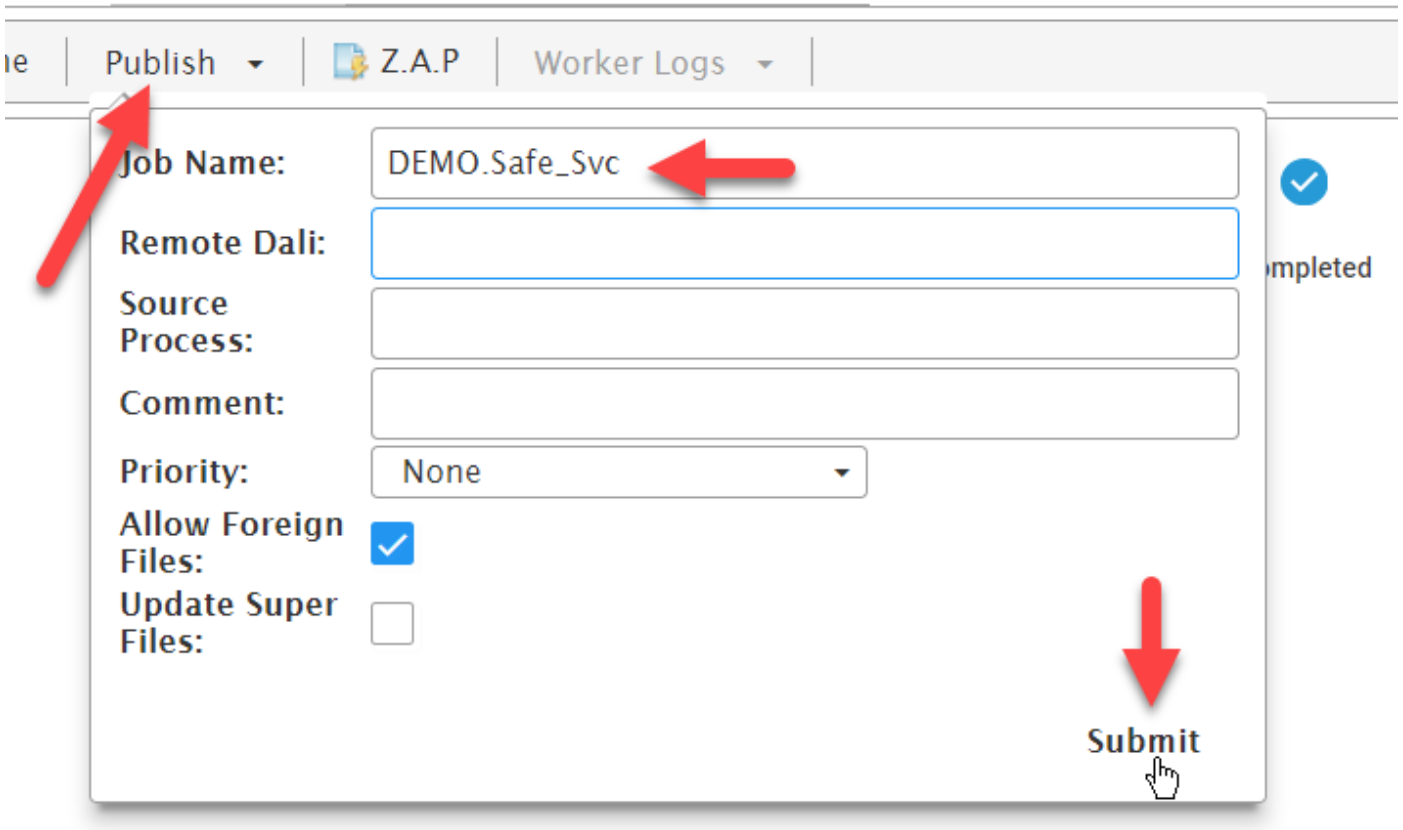
Step 3 - Deploy(Publish) and then Test Your Query:





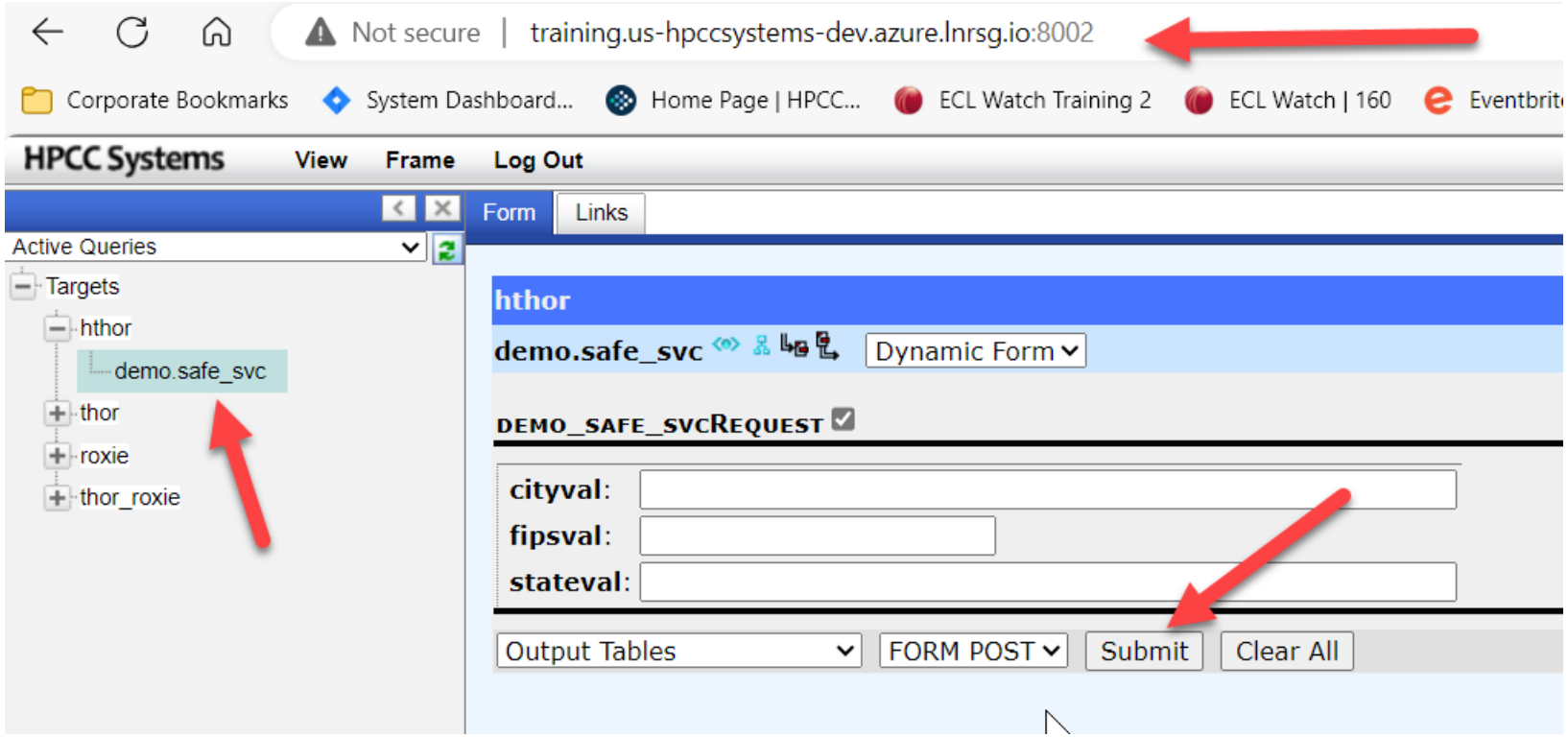
Delivering your Results!

Step 3 - Deploy(Publish) and then Test Your Query:



Delivering your Results!

Step 3 - Deploy(Publish) and then Test Your Query:



Delivering your Results!

Step 3 - Deploy(Publish) and then Test Your Query:

demo.safe svc Response

Dataset: Result 1

	city	state	name	street	zip	affiliation	primaryfips
1	ATLANTA	GA	100 PEOPLE OF FAITH ATLANTA GEORGIAINC	1949 MERCEDES CT NE	30345	3	13121
2	ATLANTA	GA	4 POINTES CHURCH OF ATLANTA INCORPORATED	1151 HAMMOND DR NE STE 240	30346	3	13121
3	ATLANTA	GA	A J FREEMAN JR MINISTRIES	PO BOX 50547	30302	3	13121
4	ATLANTA	GA	ABUNDANT GRACE COMMUNITY CHURCH OF ATLANTA INC	743 VIRGINIA AVE NE	30306	3	13121
5	ATLANTA	GA	ABUNDANT LIFE CHURCH OF GOD IN CHRIST	79 MCDDONGH BLVD SE	30315	9	13121
6	ATLANTA	GA	ABUNDANT LIFE CHURCH OF GOD IN CHRIST	PO BOX 17596	30316	9	13121
7	ATLANTA	GA	ABUNDANT LOVE UNITARIAN UNIVERSALIST	PO BOX 11372	30310	3	13121
8	ATLANTA	GA	ACTION MINISTRIES INC	1700 CENTURY CIRCLE NE NO 200	30345	3	13121
9	ATLANTA	GA	ADELPHOS FELLOWSHIP CHURCH INC	310 SPRINGDALE DR NE	30305	3	13121
10	ATLANTA	GA	AGAPE CHRISTIAN TRAINING CENTER	215 LAKEWOOD WAY SW	30315	3	13121
11	ATLANTA	GA	AGAPE DELIVERANCE MINISTRIES	791 MAGNOLIA WAY NW APT 1332	30314	3	13121
12	ATLANTA	GA	AGAPE TEMPLE WHERE JESUS BREAKS EVERY FETTER OF ATLANTA GEORGIA IN	845 HARWELL ST NW	30314	3	13121
13	ATLANTA	GA	AHAVA EARLY LEARNING CENTER INC	600 PEACHTREE BATTLE AVE NW	30327	3	13121
14	ATLANTA	GA	ALL NATIONS UNITED GOSPEL	751 FAIRBURN RD SW APT 5221	30331	3	13121
15	ATLANTA	GA	ALL NEEDS MET MINISTRIES INC	PO BOX 491708	30349	3	13121
16	ATLANTA	GA	ALL SAINTS ANGLICAN CHURCH ATLANTA	PO BOX 366491	30336	3	13121
17	ATLANTA	GA	ALLIANCE FOR CHRISTIAN MEDIA	2715 PEACHTREE RD NE	30305	3	13121
18	ATLANTA	GA	ALLIANCE OF DIVINE LOVE INC	2060 OLD GEORGIAN TER NW	30318	9	13121
19	ATLANTA	GA	ALLIANCE OF DIVINE LOVE INC	3014 WHISPERING HILLS CT	30341	9	13121
20	ATLANTA	GA	ALLIANCE OF DIVINE LOVE INC	3750 PEACHTREE RD NE APT 410	30319	9	13121
21	ATLANTA	GA	ALOC INC	2936 BLUESTONE DR SW	30331	3	13121
22	ATLANTA	GA	ALPHA & OMEGA HOLISTIC INTERNATIONAL OUTREACH APOSTOLIC	369 MCDANIEL ST SW APT 2711	30313	3	13121
23	ATLANTA	GA	ALPHARETTA CHURCH OF RELIGIOUS SCIENCE	100 HANNOVER PARK RD STE 160	30350	9	13121
24	ATLANTA	GA	AMAZED INC	1643 MT VERNON RD	30338	3	13121
25	ATLANTA	GA	AMBASSADORS FOR CHRIST NATIONAL MINISTRIES INC	2870 PHARR COURT SOUTH NW APT 710	30305	3	13121
26	ATLANTA	GA	AMERICAN ACADEMY OF RELIGION INC	825 HOUSTON MILL RD NE STE 300	30329	3	13121
27	ATLANTA	GA	AMERICAN BAPTIST HISTORICAL SOCIETY	3001 MERCER UNIVERSITY DR	30341	9	13121
28	ATLANTA	GA	ANCIENT MYSTICAL ORDER OF ROSAE CRUCIS	3600 DEKALB TECH PKY SUITE 115	30340	9	13121
29	ATLANTA	GA	ANCIENT MYSTICAL ORDER OF ROSAE CRUCIS	3600 DEKALB TECH PKY SUITE 115	30340	9	13121
30	ATLANTA	GA	ANCIENT MYSTICAL ORDER OF ROSAE CRUCIS	3600 DEKALB TECH PKY SUITE 115	30340	9	13121
31	ATLANTA	GA	ANGLICAN EPISCOPAL CHURCH OF NA INC	409 N STRATFORD RD NE	30342	3	13121

HPCC SYSTEMS

Alternate Delivery: Visualization

HPCC Systems provides built-in Visualization of your output data in a variety of charts and graphs. You can visualize your data in three ways:

- Using the Chart Tool in the ECL Playground.
- Accessing the Visualize tab in all ECL workunits
- Using the Resources tab in conjunction with the ECL Visualizer bundle.

Installing:
ecl bundle install https://github.com/hpcc-systems/Visualizer.git

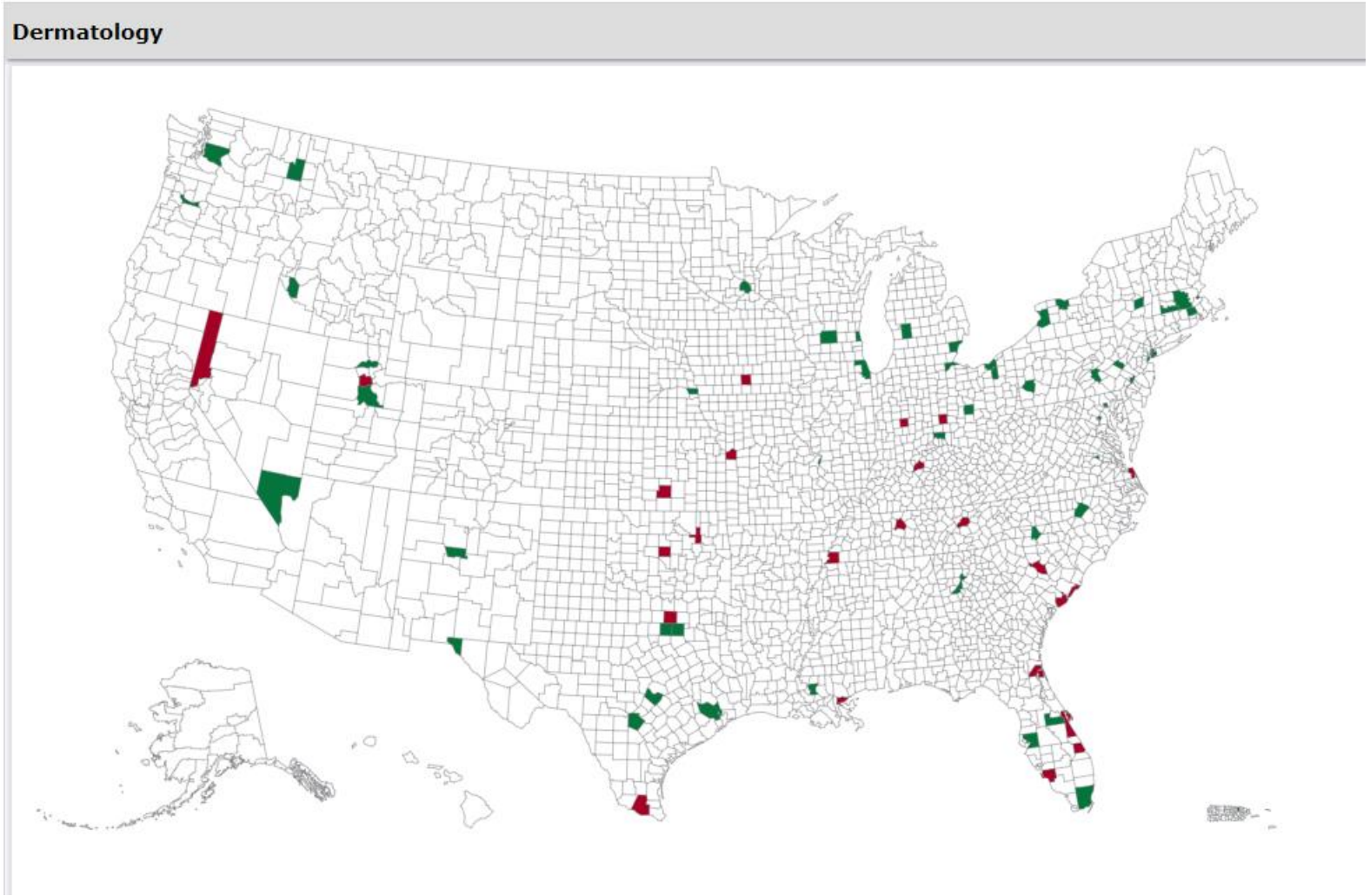
https://hpccsystems.com/resources/visualizing-ecl-and-sharing-your-results-the-hpcc-systems-visualizer/

https://github.com/hpcc-systems/Visualizer

Visualization Examples:

```
1   IMPORT $,Visualizer;
2
3   Cities := $.File_AllData.City_DS;
4
5   //Build Table
6   DensityTbl := TABLE(Cities,{(INTEGER)county_fips,(INTEGER)density});
7
8   OUTPUT(DensityTbl,NAMED('DenFIPS'));
9
10  Visualizer.Choropleth.USCounties('Fips_demo',,'DenFIPS', , , DATASET([{'paletteID', 'Defaul;t'}], Visualizer.KeyValueDef));
11
12
```

Visualization Examples:



Final Thoughts

- ✓ Since your solution is the key part to this challenge you can use `#OPTION('obfuscateOutput', TRUE);` at the start of your code to hide it from being viewed on ECL Watchpage. If you decide to use #OPTION make sure to remove if from the WUID that you shared with the judges. When obfuscateOutput set to true, details are removed from the generated workunit, including ECL code, estimates of record size, and number of records.
- ✓ If you want to write the result to a file, make sure the file name starts with your team's name for uniqueness purpose.
- ✓ Make sure the query names are unique and easy to identify. Do not use generic names like test, mentors, or roxie. We suggest adding your team's name as well. General names will result in other teams overwriting your files, queries, and results
- ✓ We encourage team play so teams that help answer questions in Slack will be considered favorably in judging.
- ✓ Direct emails and direct messages to judges asking for support will be *ignored* and it won't work in team's favor
- ✓ We also encourage students to leverage our community forum and/or StackOverflow for ECL coding related questions. Please make sure to tag your questions with *hpcc-ecl*.

HPCC SYSTEMS

Useful links!

NUhacks'24 HPCC Systems Wiki Page:
https://hpccsystems.atlassian.net/wiki/spaces/hpcc/pages/115048451/Northeastern+University+Hackathon+2024

"Learn ECL" Web Tutorial:
https://solutionslab.hpccsystems.com/learn-ecl/introduction/

ECL training containing six short videos
https://www.youtube.com/watch?time_continue=192&v=Lk78BCCtM-0

ECL documentation
http://cdn.hpccsystems.com/releases/CE-Candidate-9.8.22/docs/EN_US/ECLLanguageReference_EN_US-9.8.22-1.pdf

Visualization document
https://cdn.hpccsystems.com/releases/CE-Candidate-9.8.22/docs/EN_US/VisualizingECL_EN_US-9.8.22-1.pdf

Standard Library
https://cdn.hpccsystems.com/releases/CE-Candidate-9.8.22/docs/EN_US/ECLStandardLibraryReference_EN_US-9.8.22-1.pdf

Machine Learning
https://hpccsystems.com/download/free-modules/machine-learning-library

HPCC SYSTEMS

Good Luck to all Participants!

HPCC SYSTEMS